



I Pw

Attorney Docket No. YOR920040039US1

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent Application

Applicant(s): C.C. Aggarwal et al.  
Docket No.: YOR920040039US1  
Serial No.: 10/801,420  
Filing Date: March 16, 2004  
Group: 2171  
Examiner: To Be Assigned

I hereby certify that this paper is being deposited on this date with the U.S. Postal Service as first class mail addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Signature:  Date: August 18, 2004

Title: Methods and Apparatus for Data Stream  
- Clustering for Abnormality Monitoring

---

INFORMATION DISCLOSURE STATEMENT

Mail Stop Amendment  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Sir:

Pursuant to 37 C.F.R. §§1.56, 1.97 and 1.98, Applicants' attorney wishes to bring to the attention of the Patent and Trademark Office the following documents listed on the accompanying Form PTO-1449. A copy of each listed document is enclosed.

1. C.C. Aggarwal, "A Framework for Diagnosing Changes in Evolving Data Streams," ACM SIGMOD Conference, pp. 575-586, June 2003.
2. B. Babcock et al., "Models and Issues in Data Stream Systems," ACM PODS Conference, pp. 1-30, 2002.
3. P. Domingos et al., "Mining High-Speed Data Streams," ACM SIGKDD Conference, 10 pages, 1998.
4. S. Guha et al., "ROCK: A Robust Clustering Algorithm for Categorical Attributes," Proceedings of the International Conference on Data Engineering, 10 pages, March 1999.

5. L. O'Callaghan et al., "Streaming-Data Algorithms for High-Quality Clustering," ICDE Conference, pp. 1-25, October 2001.
6. P.S. Bradley et al., "Scaling Clustering Algorithms to Large Databases," SIGKDD Conference, 7 pages, 1998.
7. S. Guha et al., "CURE: An Efficient Clustering Algorithm for Large Databases," ACM SIGMOD Conference, pp. 73-84, 1998.
8. R.T. Ng et al., "Efficient and Effective Clustering Methods for Spatial Data Mining," Proceedings of the 20th VLDB Conference, pp. 144-155, 1994.
9. E. Knorr et al., "Algorithms for Mining Distance-Based Outliers in Large Datasets," Proceedings of the 24th VLDB Conference, 12 pages, 1998.
10. E. Knorr et al., "Finding Intensional Knowledge of Distance-Based Outliers," Proceedings of the 25th VLDB Conference, pp. 211-222, 1999.
11. S. Ramaswamy et al., "Efficient Algorithms for Mining Outliers from Large Data Sets," Proceedings of the ACM SIGMOD Conference, 12 pages, 2000.
12. T. Zhang et al., "BIRCH: An Efficient Data Clustering Method for Very Large Databases," ACM SIGMOD Conference, pp. 103-114, 1996.
13. V. Ganti et al., "CACTUS-Clustering Categorical Data Using Summaries," Proceedings of the ACM SIGKDD Conference, 11 pages, 1999.
14. D. Gibson et al., "Clustering Categorical Data: An Approach Based on Dynamical Systems," Proceedings of the VLDB Conference, pp. 1-22, 1998.
15. C.C. Aggarwal et al., "A Framework for Clustering Evolving Data Streams," Proceedings of the 29th VLDB Conference, 12 pages, 2003.

It is believed that there is no fee due in conjunction with the filing of this Information Disclosure Statement. In the event of non-payment or improper payment of a required fee, the Commissioner is authorized to charge or to credit **International Business Machines Corporation Deposit Account No. 50-0510** as required to correct the error.

Attorney Docket No. YOR920040039US1

The filing of this Information Disclosure Statement shall not be construed as a representation that a search has been made, or as an admission that the information cited is considered to be material to patentability, or as a representation that no other material information exists.

Respectfully submitted,



Date: August 18, 2004

Robert W. Griffith  
Reg. No. 48,956  
Attorney for Applicant(s)  
Ryan, Mason & Lewis, LLP  
90 Forest Avenue  
Locust Valley, NY 11560  
(516) 759-4547

**FORM PTO-1449 (MODIFIED)****LIST OF PUBLICATIONS FOR  
APPLICANT'S INFORMATION  
DISCLOSURE STATEMENT**

Applicant(s): C.C. Aggarwal et al.  
 Docket No.: YOR920040039US1  
 Serial No.: 10/801,420  
 Filing Date: March 16, 2004  
 Group: 2171

**U.S. PATENT DOCUMENTS**

EXAMINER	INITIAL	DOCUMENT NO.	DATE	NAME	FILING DATE
----------	---------	--------------	------	------	-------------

**FOREIGN PATENT DOCUMENTS**

EXAMINER	INITIAL	DOCUMENT NO.	DATE	COUNTRY	CLASS/SUBCLASS	TRANSLATION YES      NO
----------	---------	--------------	------	---------	----------------	----------------------------

**OTHER DOCUMENTS**

EXAMINER	INITIAL	REF NO.	AUTHOR, TITLE, DATE, PERTINENT PAGES, ETC.
----------	---------	---------	--

- 1. C.C. Aggarwal, "A Framework for Diagnosing Changes in Evolving Data Streams," ACM SIGMOD Conference, pp. 575-586, June 2003.
- 2. B. Babcock et al., "Models and Issues in Data Stream Systems," ACM PODS Conference, pp. 1-30, 2002.
- 3. P. Domingos et al., "Mining High-Speed Data Streams," ACM SIGKDD Conference, 10 pages, 1998.
- 4. S. Guha et al., "ROCK: A Robust Clustering Algorithm for Categorical Attributes," Proceedings of the International Conference on Data Engineering, 10 pages, March 1999.
- 5. L. O'Callaghan et al., "Streaming-Data Algorithms for High-Quality Clustering," ICDE Conference, pp. 1-25, October 2001.
- 6. P.S. Bradley et al., "Scaling Clustering Algorithms to Large Databases," SIGKDD Conference, 7 pages, 1998.
- 7. S. Guha et al., "CURE: An Efficient Clustering Algorithm for Large Databases," ACM SIGMOD Conference, pp. 73-84, 1998.
- 8. R.T. Ng et al., "Efficient and Effective Clustering Methods for Spatial Data Mining," Proceedings of the 20th VLDB Conference, pp. 144-155, 1994.
- 9. E. Knorr et al., "Algorithms for Mining Distance-Based Outliers in Large Datasets," Proceedings of the 24th VLDB Conference, 12 pages, 1998.
- 10. E. Knorr et al., "Finding Intensional Knowledge of Distance-Based Outliers," Proceedings of the 25th VLDB Conference, pp. 211-222, 1999.

Examiner

Date Considered

**Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.

**FORM PTO-1449 (MODIFIED)****LIST OF PUBLICATIONS FOR  
APPLICANT'S INFORMATION  
DISCLOSURE STATEMENT**

Applicant(s): C.C. Aggarwal et al.  
Docket No.: Y0R920040039US1  
Serial No.: 10/801,420  
Filing Date: March 16, 2004  
Group: 2171

**OTHER DOCUMENTS (cont'd.)****EXAMINER****INITIAL****REF NO.****AUTHOR, TITLE, DATE, PERTINENT PAGES, ETC.**

---

- 11. S. Ramaswamy et al., "Efficient Algorithms for Mining Outliers from Large Data Sets," Proceedings of the ACM SIGMOD Conference, 12 pages, 2000.
- 12. T. Zhang et al., "BIRCH: An Efficient Data Clustering Method for Very Large Databases," ACM SIGMOD Conference, pp. 103-114, 1996.
- 13. V. Ganti et al., "CACTUS-Clustering Categorical Data Using Summaries," Proceedings of the ACM SIGKDD Conference, 11 pages, 1999.
- 14. D. Gibson et al., "Clustering Categorical Data: An Approach Based on Dynamical Systems," Proceedings of the VLDB Conference, pp. 1-22, 1998.
- 15. C.C. Aggarwal et al., "A Framework for Clustering Evolving Data Streams," Proceedings of the 29th VLDB Conference, 12 pages, 2003.

---

**Examiner****Date Considered**

---

**Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609; draw line through citation if not in conformance and not considered. Include copy of this form with next communication to Applicant.